APPENDIX A. DATA RESOURCES ON IT IN THE HOME

THE DATASET INVENTORY

The dataset inventory was assembled from a variety of sources, primarily the World Wide Web (WWW), since most IT-related resources tend to have a substantial Web presence. In addition, several staff members of the Graphics, Visualization, and Usability (GVU) Center at Georgia Tech, who are responsible for GVU's WWW User Survey, were interviewed. GVU has researched and placed links on their Web page to many of the survey and data collection activities about computers and the Internet.

AVAILABILITY OF DATA RESOURCES

There are few data resources that meet the standards established for inclusion. To be included, data resources had to have a primary focus on technology acquisition and use by individuals or households and meet three criteria:

1. Data on information technology in the home had to be obtained through valid research methods.

2. Data had to be obtained from samples of the entire U.S. population.

3. More than 1 year of comparable data had to be available.

Eight data resources meet these criteria.[1]

At the outset, 20 potential sources of data on IT in the home were identified. Virtually all proved nonviable, however. Most were excluded because they did not provide—or no longer provided—data related to the scope of this project. Scholarly studies based on empirical research (typically one-time studies) were not included among the data resources, but are covered in the annotated works in Appendix B to this report.

Six organizations provide the data resources identified in the data inventory. The six organizations are:

- CommerceNet (commercial),
- Cyber Dialogue (commercial),
- International Data Corporation (commercial),
- Graphics, Visualization, and Usability Center at Georgia Tech (academic),
- The Pew Research Center for the People and the Press (not-for-profit), and
- U.S. Bureau of the Census (Federal Government).

Data from the last three organizations are available for free or for a nominal charge. The other organizations are commercial, for-profit market research firms.[2] This commercial dimension has implications both regarding the accessibility of data by scholars and other analysts (since prices can be as much as $10,000 for a single year's worth of data), and the ability to disseminate these data in the public domain. Because the commercial data are proprietary, their use is restricted even for subscribers.

TYPES OF DATA

Few data, if any, on the actual *impacts* of IT on the home are available. The metrics existing reflect diffusion indicators and are related to computer ownership/usage and Internet access/usage. Each data resource has a set of keywords included in its record that reflect the types of data available. These keywords and a description of each are presented in table A-1.

Note that many of the surveys from which the data are obtained differentiate between *business* and *personal* use of the computer/Internet but not between *home* use and use at *other* locations (work, school, library, etc.). Consequently, some data resources will not support analysis strictly limited to the use of IT in the home itself.

---

[1]One data resource was excluded because it focused on an individual state. This was the Research Institute for Telecommunications and Information Marketing (RITIM) at the University of Rhode Island. The RITIM data were based on a one-time survey of 1,500 households in Rhode Island. Study descriptions suggest that these data are similar to national data obtained by other organizations. For more information on the RITIM study, see http://ritim.cba.uri.edu/risurvey.

[2]Three other commercial research firms are frequently mentioned as sources of data on IT in the home: the NPD Group <http://www.npd.com>, PC Meter, and Media Metrix <http://www.mediametrix.com>. These organizations are not included here as data resources because no descriptions of the data and methods for their IT market research could be located. Also, the data product descriptions from these organizations suggest that their commercial packages reflect audience behaviors at specific websites and individual clickstream patterns. PC Meter merged with Media Metrix and is a data collection service that meters PC and Internet use much the way Nielsen meters television usage. A description of PC Meter may be found in Coffey and Stipp (1997).

THE INTERNET META-SITES

For analysts interested in statistics, data, and trends related to the use of the Internet, two Internet meta-sites offer a good starting point. These Web pages contain dozens of summaries and links to reports and studies of the Internet. The organizations that sponsor these sites do not conduct primary research of their own; rather, they serve as resources for information about the Internet. The two sites are:

- CyberAtlas http://www.cyberatlas.com, and
- Nua Internet Surveys http://www.nua.ie/surveys/index.cgi.

The remainder of this appendix consists of fact sheets summarizing the 8 datasets and 2 meta-sites.

Table A1. IT in the home data resources: keyword glossary	
Keyword	Description
General diffusion indicators	
Computer ownership..............	Computer ownership and access, technological features of home computers
Computer usage.....................	Who uses computer, how frequently, how long, for what purposes
Demographics.........................	Age, sex, ethnicity, income, education level, geographic region of users
Internet access.......................	Access and location of access to the Internet
Internet usage........................	Who uses Internet, how frequently, how long, for what purposes
Subjects addressed	
Business usage......................	Types of business use for home businesses and home offices (accounting, finance, etc.)
Citizenship..............................	Types of use related to citizenship activities: political speech and communication, use of Internet to gather political information, use of Internet to interact with government agencies, etc.
Culture & values.....................	Attitudes toward role of the Internet on relationships, values, etc.
Education...............................	Use for formal educational purposes: homework, continuing education, distance learning, degree programs, etc.
E-mail.....................................	Multiple dimensions of e-mail use (to whom, how often, etc.)
Employment & work................	Use for employment and work purposes: job hunting, telecommuting, etc.
Entertainment.........................	Use for games, travel, leisure, virtual museums, etc.
Health.....................................	Use for personal health and medical purposes
Legal & insurance...................	Use for personal legal and insurance purposes
Other......................................	Use of chatrooms, telnet, ftp, and other types of Internet activities
Personal finance.....................	Use for personal banking, credit, and finance purposes
Privacy...................................	Activities and attitudes related to information privacy and security
Shopping................................	Use for personal shopping on the Internet

CommerceNet/Nielsen Internet Demographic Survey

Abstract: Survey focus is explicitly and exclusively on uses of the Internet. Data are obtained semiannually through telephone surveys. Over 100 distinct questionnaire items explore Internet access and usage. Internet usage data cannot necessarily be isolated to the home; questions related to e-mail, personal business and finance, and other on-line activities do not always differentiate between home, work, or other access locations. CommerceNet is a commercial, for-profit market research organization.

Time Period: Survey conducted semiannually since 1995.

Availability: Data are available to subscribers through an on-line database called Gideon http://www.commerce.net/research/gideon. Most metrics are available as population count and index. Data are subscribed to in packages, which cost from $1,500 for a small subset of demographic variables to $10,000 for the complete dataset. Subscription prices are for a single-year's data.

Keywords: Computer ownership/Demographics/Internet access/Internet usage/Business usage/Citizenship/E-mail/Health/Legal & insurance/Other/Personal finance/Privacy/Shopping

URL: http://www.commerce.net/research/

Methodology: *Purpose of Data Collection:* To assess the behaviors, usage patterns, and perceptions of the Internet by individuals over the age of 16 in the United States and Canada.

Survey Method: Random digit dial telephone survey with replacement sampling for nonresponse.

Survey Design: The target population is Americans and Canadians aged 16 and over. Respondents are selected from a geographically stratified, unrestricted set of random digit frames of phone exchanges in the United States and Canada. For the June 1998 survey, approximately 175,000 phone numbers were selected. Approximately 5,000 individuals over age 16 answered the survey in June 1998; 8,000 such interviews were conducted in September 1997. Respondent weights are created to adjust for undercoverage and nonresponse. Data are adjusted by age, sex, region, education, race, and Hispanic origin to be statistically representative of the overall population.

Variables Collected: More than 100 questionnaire items elicit data on user demographics (sex, age, income, education, race, ethnicity, occupation) and Internet usage patterns (frequency of use, duration of use, types of use—e-mail, chatrooms, shopping, ftp, telnet).

Statistical Reliability: Confidence intervals/margins of error not readily available.

More detail on survey methodology may be found at http://www.commerce.net/research/gideon/index.html.

Notes: (1) The text of survey questions is available on-line at http://www.commerce.net/research/ gideon/quesmenu.htm.

(2) Interactive cross-tabulations are possible on-line.

Contact Info: CommerceNet
4005 Miranda Avenue, Suite 175
Palo Alto, CA 94304
Phone: (650) 858-1930
Fax: (650) 858-1936

CYBER DIALOGUE AMERICAN INTERNET USER SURVEY

Abstract: Cyber Dialogue has been conducting the semiannual American Internet User Survey since 1994. No detailed data on the methodology or the nature of the survey questions is readily available. Cyber Dialogue is a commercial, for-profit market research organization.

Time Period: Survey conducted semiannually since 1994.

Availability: Data are available as a stand-alone survey or as a continuous reporting service. Prices are $1,550 per year for stand-alone reports; $2,600 per year for continuous reporting services.

Keywords: Demographics/Internet access/Internet usage

URL: http://www.cyberdialogue.com/index_4.html

Methodology: *Purpose of Data Collection:* To assess the characteristics and behaviors of U.S. Internet users and potential users.

Survey Method: Random digit dial telephone survey with replacement sampling for nonresponse.

Survey Design: The target population is Americans age 18 and over. The American Internet User Survey has two design elements. The first is a random digit dial telephone survey in which households are screened on their Internet use. The screening survey is used to estimate the adult Internet user population in the United States. For the screening survey, respondents are selected from a geographically stratified set of random digit frames of phone exchanges. For the April 1997 survey, approximately 59,000 phone calls were made; approximately 13,000 households were interviewed regarding Internet use by all individuals in the household.

The second design element is in-depth interviews on Internet use with qualified individuals age 18 and over. One thousand in-depth surveys are conducted with qualified individuals who use the Internet, and 1,000 interviews are conducted with individuals who do not use the Internet. The sample of qualified Internet users is drawn from the screening interviews. To qualify for the in-depth interview, individuals had to be at least 18 years of age and use as least one Internet application other than e-mail. Data are adjusted by age, sex, census region, and level of education to correct for response bias.

Variables Collected: The in-depth interviews cover more than 155 questionnaire items on user demographics (sex, age, education) and Internet usage patterns (frequency of use, duration of use, types of use—e.g., e-mail, chatrooms, shopping, ftp, telnet—etc.).

Statistical Reliability: At the 95 percent confidence level for the 1997 survey, subsets of the total sample reflect the following margins of error: subset sample size of 1,000, ±3.1 percent; subset sample size of 800, ±3.5 percent; subset sample size of 600, ±4.0 percent; subset sample size of 400, ±4.9 percent; subset sample size of 200, ±6.9 percent; and subset sample size of 100, ± 9.8 percent.

For more detail on survey methodology, see Peter Clemente, *State of the Net* (New York: McGraw-Hill, 1998).

Notes: The text of survey questions is not available on-line.

Contact Info: Cyber Dialogue
304 Hudson Street, 6th Floor
New York, NY 10013
Phone: (212) 255-6655
Fax: (212) 255-6622

CYBER DIALOGUE CYBERCITIZEN FINANCE SURVEY

Abstract: Cyber Dialogue initiated the Cybercitizen Finance Survey in 1998. This new survey is intended to track consumer demand for on-line financial services, particularly brokerage/investing, banking, credit card, and insurance services. Cyber Dialogue is a commercial, for-profit market research organization.

Time Period: Survey first fielded in June 1998.

Availability: Data from the Cybercitizen Finance Survey is available in reports, a complete data summary book, and unlimited access to the data by subscribers through an on-line browser. No price information is readily available.

Keywords: Demographics/Internet access/Internet usage/Business usage/Personal finance

URL: http://www.cyberdialogue.com/index_4.html

Methodology: *Purpose of Data Collection:* To track consumer demand for on-line financial services, particularly brokerage investing, banking, credit card, and insurance services.

Survey Method: Random digit dial telephone survey with replacement sampling for nonresponse.

Survey Design: The target population is Americans age 18 and over. Responses include data from 1,000 U.S. adult Internet users and 500 adult nonusers. More than 150 questions focus on consumer financial planning activities both on and off the Internet. Cyber Dialogue reports that "The definition of 'online use' utilized to establish the 1,000 completed user interviews required randomly identified U.S. adults aged 18 years or older to confirm that they use commercial on-line services or one or more Internet applications for business, personal, or academic purposes." No detailed data on response rates is readily available.

Variables Collected: No detailed information is readily available.

Statistical Reliability: No information is readily available.

No other information on the methodology or statistical reliability of this survey is readily available.

Notes: The text of survey questions is not available on-line.

Contact Info: Cyber Dialogue
304 Hudson Street, 6th Floor
New York, NY 10013
Phone: (212) 255-6655
Fax: (212) 255-6622

CYBER DIALOGUE CYBERCITIZEN HEALTH

Abstract: Cybercitizen Health is the newest of Cyber Dialogue's consumer-related Internet survey products. Its purpose is to track consumer behavior, attitudes, and demographics related to use of the Internet for health information. Cyber Dialogue is a commercial, for-profit market research organization.

Time Period: No detailed information readily available.

Keywords: Demographics/Internet usage/Health

URL: http://www.cyberdialogue.com/products/isg/cybercitizen_health.html

Methodology: *Purpose of Data Collection:* To track consumer behavior, attitudes, and demographics related to use of the Internet for health information. Cyber Dialogue's research agenda for this survey is to identify and size current and future markets for on-line health information, products, and services; segment on-line consumers by health-related consumption, attitudes toward health, and on-line usage and intensity; compare the effectiveness of the Internet for consumer marketing with other media; evaluate the profitability of various business models on-line, including those based on advertising, subscriptions, patient referrals, and commerce; provide a model for identifying and acquiring the most valuable on-line consumers; and understand how consumer concerns about privacy, security, and trust affect on-line marketing efforts.

Survey Method: A random sample telephone survey of more than 2,000 U.S. adult Internet users and nonusers, fielded in the second quarter of each year.

Survey Design: No additional information readily available.

Variables Collected: No detailed information readily available.

Statistical Reliability: No information readily available.

No other information on the methodology or statistical reliability of this survey is readily available.

Notes: Cyber Dialogue
304 Hudson Street, 6th Floor
New York, NY 10013
Phone: (212) 255-6655
Fax: (212) 255-6622

CYBERATLAS

Abstract: CyberAtlas is an Internet meta-site of on-line facts, statistics, and marketing information related to the Internet. CyberAtlas conducts no original survey research of its own; rather, it provides summaries of, and links to, dozens of studies and surveys on Internet access and usage.

Time Period: Not applicable.

Availability: Not applicable.

Keywords: Demographics/Internet usage/Business usage/Education/Employment & work/Entertainment/ Other/Personal finance/Shopping

URL: http://www.cyberatlas.com

Methodology: Not applicable.

Notes: None.

Contact Info: Not applicable.

GRAPHICS, VISUALIZATION, AND USABILITY (GVU) CENTER, GEORGIA TECH WORLD WIDE WEB (WWW) USER SURVEY

Abstract: The GVU Center has been conducting semiannual Internet use surveys since 1994. The surveys are nonrandom participant surveys completed on the Internet itself. Comparison of GVU WWW User Survey results to Internet user surveys that utilize random techniques reveals that the main area where GVU surveys show a bias exists in the experience, intensity of usage, and skill sets of the users, but not the core demographics of users. Data are freely available on the Web as a public service, although they are not highly detailed with respect to IT and the home. The GVU Center is a research center at the Georgia Institute of Technology.

Time Period: Survey conducted semiannually since 1994.

Availability: Data from each survey are freely available on-line.

Keywords: Demographics/Internet usage/Culture & values/Privacy/Shopping

URL: http://www.cc.gatech.edu/gvu/user_surveys/

Methodology: *Purpose of Data Collection:* To profile Internet users, their characteristics, and why they use the Web so as to improve development of Web-related tools and technologies and to make the Web more usable by all users.

Survey Method: The WWW User Survey is a voluntary, Web-based questionnaire that reflects a nonprobabilistic sample (i.e., survey participants are not randomly selected from a defined population of Internet users).

Survey Design: The target population is all Internet users throughout the world. Participants are solicited through several Web-based means, including announcements on Internet related newsgroups (e.g. comp.infosystems.www.announce, comp.internet.net-happenings, etc.); banners randomly rotated though high-exposure sites (e.g. Yahoo, CNN, Excite, Webcrawler, etc.); banners rotated through advertising networks (e.g. DoubleClick); announcements made to the WWW-surveying mailing list, a list maintained by GVU's WWW User Surveys composed of people interested in the surveys; and, announcements made in the popular media, (e.g., newspapers, trade magazines, etc.). Participants are thus self-selected and voluntarily complete the on-line survey. The number of participants varies from 1,500 in the first GVU survey in 1994 to more than 5,000 respondents for the tenth survey in 1998.

Variables Collected: Highly detailed demographic data: age; income; race/ethnicity; employment status; occupation; marital status; physical impairment; sex; education; location (country); location (rural, urban, suburban); frequency and duration of Internet use; types of Internet use; attitudes toward the Internet.

Statistical Reliability: Confidence intervals and margins of error do not apply since this is a nonprobabilistic sample.

Data Limitations: The GVU Center reports that "comparisons of the GVU's WWW User Survey results to other published WWW User data that utilize random techniques reveal that the main area where GVU's surveys show a bias exists in the experience, intensity of usage, and

skill sets of the users, but not the core demographics of users. Given the limitations that exist in the data as a result of the methodology, we make the following recommendation to those using the data presented within this report: We recommend that the GVU data be used with the understanding that the data has a bias towards the experienced and more frequent users than random digit dial surveys. We recommend that users interested in understanding the complete spectrum of the Internet and WWW communities augment the GVU data with random sample surveys."

Additional detail on survey methodology is at <http://www.gvu.gatech.edu/user_surveys/survey-1998-10/#methodology>.

Notes: The text of survey questions is available on-line.

Contact Info: GVU Center
College of Computing
Georgia Institute of Technology
Atlanta, GA 30332-0280
email: www-survey@cc.gatech.edu

INTERNATIONAL DATA CORPORATION (IDC)
WORK-AT-HOME SURVEY

Abstract: Survey provides key information on the changing nature of working from home, and focuses on technology acquisition and use, key attitudes, and demographic variables. Major areas of coverage include home businesses, telecommuting, and other forms of corporate home working. IDC is a for-profit commercial marketing organization.

Time Period: Survey conducted annually since 1986.

Availability: Data are available in commercial reports issued through IDC's Home Office Market Program. Reports range in price from $3,000 to $5,000 for a single year of data.

Keywords: Computer ownership/Computer usage/Demographics/Internet access/Internet usage/Business usage/Employment & work

URL: http://www.idc.com/idc7/default.htm

Methodology: Random digit dial telephone survey of 2,000 U.S. households. No other information on survey methodology or on the statistical reliability of the data is readily available.

Notes: The text of survey questions is not available on-line.

Contact Info: IDC Corporate Headquarters
Five Speen Street
Framingham, MA 01701
Phone: 508-872-8200

NUA INTERNET SURVEYS

Abstract: Internet meta-site of on-line facts, statistics, and marketing information related to the Internet. Nua conducts no original survey research of its own; rather, it provides summaries of, and links to, dozens of studies and surveys on Internet access and usage.

Time Period: Not applicable.

Availability: Not applicable.

Keywords: Demographics/Internet usage/Business usage/Employment & work/Entertainment/Other/Personal Finance/Shopping

URL: http://www.nua.ie/surveys/index.cgi

Methodology: Not applicable.

Notes: None.

Contact Info: Not applicable.

THE PEW RESEARCH CENTER FOR THE PEOPLE AND THE PRESS 1998 TECHNOLOGY SURVEY

Abstract: The Pew Research Center for the People and the Press is a nonprofit organization which has been conducting research on technology in the American home since 1994. Large-scale surveys that measure attitudes toward and use of electronic IT have been administered periodically over the past several years. Data are obtained through random digit dial telephone surveys, and tabulations for the most recent (1998) Technology Survey contain historical data for each question. This is a rich source of information on IT and the home, particularly as it relates to the *actual use* of IT.

Time Period: 1994–98

Availability: Data tables are freely available on the Web.

Keywords: Computer ownership/Computer usage/Demographics/Internet access/Internet usage/Citizenship/Culture & values/Education/E-mail/Employment & work/Entertainment/Personal Finance/Privacy/Shopping

URL: http://www.people-press.org/tech98que.htm

Methodology: *Purpose of Data Collection:* To measure attitudes toward and use of electronic IT in the home.

Survey Method: Random digit dial telephone survey with replacement sampling for nonresponse.

Survey Design: The target population is all individuals 18 years or older residing in the continental United States. The sample for this survey is based on a geographically stratified, random digit dial of working banks of telephone numbers in the continental United States. Results for the 1998 Technology Survey are based on telephone interviews conducted with 3,184 adults in the last quarter of 1998. The survey included interviews with a national sample of all adults, and an oversample of 1,184 adults who use the Internet. Results were also drawn from a survey of a nationwide sample of 1,201 adults conducted in December 1998. Respondent weights are created to adjust for nonresponse. Demographic weighting parameters are used to adjust the data to make responses representative of U.S. households with adults 18 years or older. Data on sample size and response rates are not readily available.

Variables Collected: Highly detailed questions/response categories on computer and Internet access, use, and types of use. Limited demographic data are presented; only educational status, employment status, and political party affiliation are shown in the data.

Statistical Reliability: For survey results based on the general public (N = 2,000), the confidence level is 95 percent with error attributable to sampling and other random effects of ±3 percentage points. For results based on Internet users only (N = 1,993) or Internet users who go on-line for news (N = 1,390), sampling error is ±3 percentage points. For results based on Internet users who go on-line for election news (N = 315), sampling error is ±6 percentage points. For results based on subsamples of either the general public or Internet users, sampling error is ±3.5 percentage points. Note that in addition to sampling error, question wording and logistical difficulties in conducting surveys can also introduce error or bias into the findings.

More detail on survey methodology is at http://www.people-press.org/tech98que.htm.

Notes: (1) The text of survey questions is available on-line at http://www.people-press.org/tech98que.htm.

(2) The Pew Research Center has five key research themes, one of which is Technology in the American Home, which focuses on IT. Related survey questions appear in many of the dozens of surveys the center conducts, so it is difficult to isolate a single survey series that contains appropriate data. The data tables for the 1998 Technology Survey contain data for all previous times that an individual question may have been asked, so to this extent the reported findings contain comparable historical data.

Contact Info: The Pew Research Center for the People and the Press
1150 18th Street, NW
Suite 975
Washington, DC 20036
Phone: (202) 293-3126
Fax: (202) 293-2569

U.S. BUREAU OF THE CENSUS CURRENT POPULATION SURVEY (CPS)

Abstract: The Current Population Survey is conducted monthly by the U.S. Bureau of the Census. Every few years, questions related to computer ownership, computer usage, Internet access, and Internet usage are included in the monthly survey. These data are among the few resources that clearly distinguish home use of computers and the Internet from other locations of use.

Time Period: October 1984; October 1989; October 1993; November 1994; October 1997; December 1998.

Availability: Reports and tabulations for October 1984, 1989, 1993, and 1997 are available at http://www.census.gov/population/www/socdemo/computer.html. Tabulations for November 1994, October 1997, and December 1998 are available at www.bls.census.gov/cps/computer/computer.htm.

Keywords: Computer ownership/Computer usage/Demographics/Internet access/Internet usage/Business usage/Education/E-mail/Employment & work/Entertainment/Other/Personal finance/Shopping

URLs: http:/www.census.gov/population/www/socdemo/computer.html
http://www.bls.census.gov/cps/computer/computer.htm

Methodology: *Purpose of Data Collection:* To collect detailed data on labor force characteristics of the U.S. population. The monthly CPS is occasionally supplemented with specialized questions related to topics of national interest such as health care, school enrollment, and computer usage.

Survey Method and Design: Data are based on interviews with a randomly selected sample of 50,000 U.S. housing units, with about a 94-percent response rate. The present CPS sample covers all 50 states and the District of Columbia.

Variables Collected: There are several special supplements to the CPS that are IT-related: these were conducted as part of the CPS in October 1993, November 1994, and October 1997. Additional data are available for the years 1984 and 1989; these earlier data are reported in the tabulations for the October 1993 supplement. The November 1994 supplement contains, in addition to the highly detailed demographic data, data on types of computer use in the home. The 1997, 1998, and 2000 supplements include questions related to home computer and Internet access and use.

Statistical Reliability: Standard errors and confidence intervals are reported at 90 percent confidence levels and .10 levels of significance. Detailed parameter estimates for the 1994-1998 supplements may be found at www.bls.census.gov/cps/computer/computer.htm.

More detailed information on the CPS methodology in general and the computer surveys since 1994 in particular may be found at http://www.bls.census.gov/cps/mdocmain.htm.

Notes: (1) October 1993 data tabulations (including 1989 and 1984 historical comparisons) are available at http://www.census.gov/population/www/socdemo/computer.html.

(2) The 1994 Computer Uses data and the 1997 Internet Usage data are located at http://www.bls.census.gov/cps/computer/computer.htm.

(3) Interactive cross-tabulations are not possible on-line.

Contact Info: Census Bureau
Customer Service Division
Phone: (301) 457-4100